

# Kernel estimation of the intensity of Cox processes

Nicolas KLUTCHNIKOFF

IRMAR, Université de Rennes 2, CNRS, UEB  
Campus Villejean (Rennes)  
Place du recteur Henri Le Moal  
CS 24307  
35043 Rennes cedex  
nicolas.klutchnikoff.univ-rennes2.fr

Gaspar MASSIOT

IRMAR, ENS Rennes, CNRS, UEB  
Campus de Ker Lann  
Avenue Robert Schuman, 35170 Bruz, France  
gaspar.massiot@univ-rennes1.fr

## 1 Introduction

Counting processes and in particular Cox processes have been used for many years to model a large variety of situations from neuroscience [see 2, 7, 10] to seismic [see 12], financial [see 11], insurance [see 1] or biophysical data [see 9]. Recall that a Cox process  $N = (N_t)_{t \in [0,1]}$  with random intensity  $\lambda = (\lambda(t))_{t \in [0,1]}$  is a counting process such that the conditional distribution of  $N$  given  $\lambda$  is a Poisson process with intensity  $\lambda$ . In all the previous situations one of the main problem can be summarized as the estimation of the intensity  $\lambda$  of the process [see 14].

Note that when Cox process data arise the intensity of the process is mainly not directly observed but a co-process is observed instead. Returning to one of the previous example, in single-molecule experiments only the peaks inducing the counting process and an underlying process are observed [see 9]. Another example can be found in car insurance [see 1] where the counting process models the occurrence of car crash that are subject to weather conditions. In these cases the counting process  $N = (N_t)_{t \in [0,1]}$  that naturally raises is accompanied with a co-process  $Z = (Z_t)_{t \in [0,1]}$  such that the conditional law of  $N$  given  $Z$  is a Poisson

process with intensity  $\theta(t, Z)$  where  $\theta$  is a deterministic function. By a slight abuse we shall call Cox process such a counting process. From a statistical point of view one of the major issue is to estimate the deterministic function  $\theta$  using  $n$  independent copies  $(N^1, Z^1), \dots, (N^n, Z^n)$  of  $(N, Z)$ . However, such an approach is subject to the curse of dimensionality as the covariate  $Z$  takes its values in an infinite dimension space as seen in O'Sullivan [13].

When dealing with practical problems it is often unnecessary, or at least not strictly required for the modeling, to observe the full trajectory of the co-process. One can instead observe the values taken by the co-process at some well chosen random times that cover most of the information in the co-process. In this model the co-process is observed at a finite number of random times thereby circumventing the curse of dimensionality.

In this paper we consider the following model: let  $N = (N_t)_{t \in [0,1]}$  be a counting process and  $Z = (Z_t)_{t \in [0,1]}$  be a  $\mathbb{R}^d$ -valued co-process. We assume that  $N$  admits a random intensity which depends on  $t$  and on the observations of  $Z$  at random times  $S_1 < S_2 < \dots$ .

More precisely, given the  $\sigma$ -algebra  $\mathcal{S}$  generated by these times,  $N$  is a Cox process with intensity

$$\lambda(t, Z) = \theta_S \left( t, \vec{Z}_S(t) \right), \quad (1.1)$$

where  $M$  is the counting process associated to  $S = (S_1, S_2, \dots)$ , for any function  $z : [0, 1] \rightarrow \mathbb{R}$ ,  $\vec{z}_S(t)$  denotes the projection  $(z_{S_1}, \dots, z_{S_{M_t}}) \in \mathbb{R}^{dM_t}$  and  $\theta_S(t, \cdot)$  is a function from  $\mathbb{R}^{dM_t}$  into  $\mathbb{R}_+$ .

In the sequel we consider that given  $\mathcal{S}$ ,  $(N^1, Z^1), \dots, (N^n, Z^n)$  are independent and identically distributed (*i.i.d.*) copies of  $(N, Z)$ . The goal of this paper is to construct and study the statistical properties of a kernel-type estimator of  $\lambda$  using these data. Note that the dimension of our estimation problem, which depends on the counting process  $(M_t)_{t \in [0,1]}$ , increases with  $t$ . This potentially leads to a deterioration of the accuracy of any estimation procedure as the time variable increases.

We consider a substantial data set of historical prices of 495 companies and the crude oil prices over a period of roughly one year and two months (from 17th April, 2014 to 23rd June, 2015). The Cox process data consist of the count of the number of times when the percent returns of said companies go below a certain threshold with the counting rate depending on the stochastic dynamics

of the company market capitalization. In this example the company market capitalization is represented by the action's trade volume normalized increments and is observed when the percent return of the crude oil action goes below another threshold. By analyzing this count, we aim to learn the financial properties of this 495 companies system.

The paper is organized as follows. Section 2 presents the estimator we propose and its asymptotic properties. In Section 3 we proceed to a simulation study. Then in Section 4 we apply the proposed estimator on the real data set presented above. Technical proofs of the asymptotic properties are postponed to Section 5.

## 2 Estimation strategy and results

### 2.1 Estimation strategy

Let  $t \in [0, 1]$  and  $z : [0, 1] \rightarrow \mathbb{R}$  be fixed. In this section we present the main ideas behind the construction of our estimator  $\tilde{\lambda}(t, z)$  of  $\lambda(t, z)$ .

As an introduction to our methodology we consider the ideal case where we observe  $\theta_S(t, \vec{Z}_S^k(t))$  for all  $k = 1, \dots, n$ . Then our problem of estimation can simply be viewed as a regression estimation problem where  $\lambda(t, \cdot)$  is the regression function. In this context, the Nadaraya-Watson estimator writes

$$\hat{\lambda}_{NW}(t, z) = \frac{\sum_{k=1}^n \theta_S(t, \vec{Z}_S^k(t)) H_\eta(\vec{z}_S(t) - \vec{Z}_S^k(t))}{\sum_{l=1}^n H_\eta(\vec{z}_S(t) - \vec{Z}_S^l(t))}.$$

where  $H_\eta$  denotes the multivariate product kernel  $\mathcal{H}_\eta^{\otimes dM_t}$  where  $\mathcal{H}$  is a kernel, that is  $\mathcal{H} \in L^1(\mathbb{R})$  such that  $\int_{\mathbb{R}} \mathcal{H}(u) du = 1$ ,  $\eta$  is an  $\mathcal{S}$ -measurable positive random variable (called a bandwidth) and  $\mathcal{H}_\eta(\cdot) = \eta^{-1} \mathcal{H}(\eta^{-1} \cdot)$ .

In practice  $\theta_S(t, \vec{Z}_S^k(t))$  can be estimated using the observations. Indeed, conditionally to  $\mathcal{S}$  and the  $\sigma$ -algebra  $\mathcal{Z}$  generated by the co-processes  $(Z^1, \dots, Z^n)$ ,  $N$  is a non-homogeneous Poisson process with intensity function  $t \mapsto \theta_S(t, \vec{Z}_S^k(t))$ , a natural estimator of this intensity is given by

$$\int_0^t K_h(t-u) dN_u^k = \sum_{i=1}^{N_t^k} K_h(t - T_i^k),$$

where  $T_1^k, T_2^k, \dots$  denote the jumping times of the trajectory  $N^k$ ,  $K : \mathbb{R}_+ \rightarrow \mathbb{R}$  is a kernel and  $h$  is a bandwidth. Denoting

$$\begin{aligned}\hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) &= \frac{1}{n} \sum_{k=1}^n \sum_{i=1}^{N_t^k} K_h(t - T_i^k) H_\eta(\vec{z}_S(t) - \vec{Z}_S^k(t)), \\ \hat{f}_{S,\eta}(\vec{z}_S(t)) &= \frac{1}{n} \sum_{l=1}^n H_\eta(\vec{z}_S(t) - \vec{Z}_S^l(t)),\end{aligned}$$

we define the plug-in estimator by

$$\hat{\lambda}(t, z) = \frac{\hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t))}{\hat{f}_{S,\eta}(\vec{z}_S(t))}.$$

For the sake of stability [see 4] we consider a trimmed version of the previous estimator

$$\tilde{\lambda}(t, z) = \tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) = \frac{\hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t))}{\tilde{f}_{S,\eta}(\vec{z}_S(t))}, \quad (2.1)$$

where  $\tilde{f}_{S,\eta}(\vec{z}_S(t)) = \hat{f}_{S,\eta}(\vec{z}_S(t)) \vee a_n$  and  $(a_n)_{n \in \mathbb{N}}$  is an  $\mathcal{S}$ -measurable real-valued positive decreasing sequence.

## 2.2 Results

In this paper we are interested in the local behaviour of our estimator. We thus consider the pointwise mean squared error defined by

$$\text{MSE}(t, z) = \mathbb{E} \left[ \left( \tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) - \theta_S(t, \vec{z}_S(t)) \right)^2 \right], \quad (2.2)$$

and we make local regularity assumptions on the model.

Remark that, almost surely,  $t \in ]S_{M_t}, S_{M_t} + 1[ = I_{M_t}$ . For any  $u \in I_{M_t}$ ,  $\vec{Z}_S(u) = \vec{Z}_S(t) \in \mathbb{R}^{dM_t}$  and  $\theta_{S,M_t}(u, \cdot) = \theta_S(u, \cdot)$  is defined from  $\mathbb{R}^{dM_t}$  into  $\mathbb{R}_+$ .

**(H1)** Given  $\mathcal{S}$ , for any  $u \in I_{M_t}$ ,  $\vec{Z}_S(u)$  admits a conditional density  $f_{M_t}$  defined from  $\mathbb{R}^{dM_t}$  into  $\mathbb{R}_+$ ;

**(H2)**  $\theta_{S,M_t}$  and  $f_{M_t}$  are positive continuous functions;

(H3)  $\theta_{S,M_t}$  and  $f_{M_t}$  are twice differentiable and there exists a random variable  $Q_{M_t}$  depending on the dimension  $dM_t$  such that for all  $1 \leq k \leq dM_t$ ,

$$\forall y \in \mathbb{R}^{dM_t}, \left\| \frac{\partial^2}{\partial x_k^2} f_{M_t}(y) \right\|_2 \leq Q_{M_t},$$

and for all  $1 \leq k \leq dM_t + 1$ ,

$$\forall (y, u) \in \mathbb{R}^{dM_t} \times I_{M_t}, \left\| \frac{\partial^2}{\partial x_k^2} \left( f_{M_t}(y) \theta_{S,M_t}(u, \vec{y}_S(u)) \right) \right\|_2 \leq Q_{M_t},$$

where  $\|\cdot\|_2$  is the euclidean norm;

(H4) There exists positive constants  $F_0$ ,  $F_\infty$  and  $\Theta$  such that  $\|\theta_{S,M_t}\|_\infty < \Theta^{M_t}$  and  $0 < F_0^{M_t} \leq f_{M_t} \leq F_\infty^{M_t} < \infty$ .

We also make technical assumptions on the kernels, the sequence  $a_n$  and  $Q_{M_t}$

(H5)  $\mathcal{H}$  is a kernel of order 2 (that is, for all  $j$  from 1 to 2,  $\int_{\mathbb{R}} u^j \mathcal{H}(u) du = 0$  and  $\int_{\mathbb{R}} |\mathcal{H}(u)| du < +\infty$ ),  $\text{supp} \mathcal{H} = [-1, 1]$  and  $\|\mathcal{H}\|_\infty < \infty$ ;

(H6)  $K$  is a kernel of order 2,  $K \in \mathbb{L}^4(\mathbb{R})$  and  $\text{supp} K = [0, 1]$ ;

(H7)  $a_n = (n\eta^{dM_t})^{\varepsilon-1}$  for some  $\varepsilon \in (0, 1/2)$ ;

(H8) There exists a positive constant  $Q_0(t)$  such that  $\mathbb{E}[M_t^8 Q_{M_t}^4] \leq Q_0(t)$ . Moreover, for any  $\lambda > 0$ , there exists a positive constant  $Q_1(\lambda, t)$  such that  $\mathbb{E}e^{\lambda M_t} < Q_1(\lambda, t)$ .

We are now in position to state our main results. Define the pointwise conditional mean squared error by

$$\text{MSE}_{\mathcal{S}}(t, z) = \mathbb{E}_{\mathcal{S}} \left[ \left( \tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) - \theta_S(t, \vec{z}_S(t)) \right)^2 \right].$$

**Theorem 1.** Assume that (H1) to (H8) are satisfied. Let  $h$  and  $\eta$  be two  $\mathcal{S}$ -measurable bandwidths such that  $h \rightarrow 0$ ,  $\eta \rightarrow 0$ ,  $nh\eta^{dM_t} \rightarrow +\infty$  and  $n\eta^{dM_t+4} \rightarrow 0$  almost surely (a.s.) as  $n \rightarrow +\infty$  then the pointwise conditional mean squared error writes

$$\text{MSE}_{\mathcal{S}}(t, z) = \underset{n \rightarrow +\infty}{O} \left( h^4 + \eta^4 + \frac{1}{nh\eta^{dM_t}} \right), \quad \mathbb{P} - a.s. \quad (2.3)$$

Note that the optimal choice of the  $\mathcal{S}$ -measurable bandwidths  $h$  and  $\eta$  for the pointwise conditional mean squared error is then  $h = \eta = n^{-\frac{1}{5+dM_t}}$ . This leads to the following corollary for the control of the pointwise mean squared error.

**Corollary 2.** *Under the assumptions of theorem 1, the pointwise mean squared error writes*

$$\text{MSE}(t, z) = \underset{n \rightarrow +\infty}{O} \left( \mathbb{E} \left( n^{-\frac{4}{5+dM_t}} \right) \right), \mathbb{P} - a.s. \quad (2.4)$$

We can get the consistency of our estimator under weaker assumptions as shown in the following proposition.

**Proposition 3.** *Assume that (H1), (H2) and (H4) to (H7) are satisfied. Let  $h$  and  $\eta$  be two  $\mathcal{S}$ -measurable bandwidths such that  $h \rightarrow 0$ ,  $\eta \rightarrow 0$ ,  $nh\eta^{dM_t} \rightarrow +\infty$  a.s. as  $n \rightarrow +\infty$ , then*

$$\tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) \xrightarrow{\mathbb{P}} \theta_S(t, \vec{z}_S(t)).$$

Theorem 4 shows the asymptotic normality of our estimator.

**Theorem 4.** *Assume that (H1) to (H8) are satisfied. Let  $h$  and  $\eta$  be two  $\mathcal{S}$ -measurable bandwidths such that  $h \rightarrow 0$ ,  $\eta \rightarrow 0$ ,  $nh\eta^{dM_t} \rightarrow +\infty$ ,  $nh\eta^{dM_t+4} \rightarrow 0$ , and  $nh^5\eta^{dM_t} \rightarrow 0$  a.s. as  $n \rightarrow +\infty$  then for any  $z : [0, 1] \rightarrow \mathbb{R}$  such that  $\theta_S(t, \vec{z}_S(t)) \neq 0$*

$$\left( nh\eta^{dM_t} \right)^{1/2} \frac{\tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) - \theta_S(t, \vec{z}_S(t))}{\left[ \tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t} / \tilde{f}_{S,\eta}(\vec{z}_S(t)) \right]^{1/2}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

*Remarks.* Corollary 2 gives the tools to define optimal bandwidths in terms of pointwise asymptotic mean squared error. Assumptions must however be made on the process  $M$  to conclude on the convergence rate since the MSE depends on the quantity  $\mathbb{E} \left( n^{-\frac{4}{5+dM_t}} \right)$ . In what follows we assume that  $M$  is a renewal process with inter-arrival times distributed accordingly to a strictly increasing cumulative distribution function  $F$ . The behaviour of the considered expectation is linked to the local behaviour of  $F$  around 0. The two following examples give incentive on the performances of our estimator for  $F$  close to 0 around 0.

**Example 1.** Let  $\varepsilon$  be a positive constant and assume that  $F(x) = 0$  for any  $x \in [0, \varepsilon]$ . Then,

$$\begin{aligned} \mathbb{E} \left( n^{-\frac{4}{5+dM_t}} \right) &\leq n^{-\frac{4}{5}} \mathbb{P}(S_0 \leq t) + \sum_{k \geq 1} \left[ n^{-\frac{4}{5+dk}} - n^{-\frac{4}{5+d(k-1)}} \right] \mathbb{P}(S_k \leq t) \\ &\leq n^{-\frac{4}{5}} + \sum_{k=1}^{\lfloor \frac{t}{\varepsilon} \rfloor} \left[ n^{-\frac{4}{5+dk}} - n^{-\frac{4}{5+d(k-1)}} \right] \\ &\leq n^{-\frac{4}{5+d \lfloor \frac{t}{\varepsilon} \rfloor}}. \end{aligned}$$

So that (2) gives

$$\text{MSE}(t, z) = O_{n \rightarrow +\infty} \left( n^{-\frac{4}{5+d \lfloor \frac{t}{\varepsilon} \rfloor}} \right).$$

This is the optimal rate of convergence for the nonparametric regression with a twice continuously differentiable regression function from  $\mathbb{R}^{d \lfloor \frac{t}{\varepsilon} \rfloor + 1}$  to  $\mathbb{R}$  [see 8].

**Example 2.** Let  $\varepsilon$  and  $\alpha$  be two positive constants such that  $\alpha > 1$  and assume that  $F(x) \leq \exp \{ -(\varepsilon x^{-1})^\alpha \}$ , around 0. Then,

$$\begin{aligned} \mathbb{E} (a^{M_t}) &\leq n^{-\frac{4}{5}} \mathbb{P}(S_0 \leq t) + \sum_{k \geq 1} \left[ n^{-\frac{4}{5+dk}} - n^{-\frac{4}{5+d(k-1)}} \right] \mathbb{P}(S_k \leq t) \\ &\leq n^{-\frac{4}{5+dk^*}} + \sum_{k \geq k^*+1} k F \left( \frac{\varepsilon}{(3 \log k)^{1/\alpha}} \right) \\ &\leq n^{-\frac{4}{5+dk^*}}, \end{aligned}$$

where  $k^* = \left( \frac{t}{\varepsilon} \right)^{\frac{\alpha}{\alpha-1}} 3^{\frac{1}{\alpha-1}}$ . So that (2) gives

$$\text{MSE}(t, z) = O_{n \rightarrow +\infty} \left( n^{-\frac{4}{5+d \left( \frac{t}{\varepsilon} \right)^{\frac{\alpha}{\alpha-1}} 3^{\frac{1}{\alpha-1}}}} \right).$$

Remark that if we formally take  $\alpha = +\infty$ , we get back to the situation of Example 1 and the upper bounds coincide as the previous upper bound writes

$$\text{MSE}(t, z) = O_{n \rightarrow +\infty} \left( n^{-\frac{4}{5+d \frac{t}{\varepsilon}}} \right).$$

Note that the rate of convergence of the mean squared error is in-between the traditional finite-dimensional rate [see 8], and the rates obtained by Biau et al [3] in the infinite-dimensional setting. This is explained by the particularity of our model which is itself in-between the finite and infinite-dimensional settings.

### 3 Simulation study

In this section we aim at studying the performances of our estimator from a practical point of view. To this end we study our estimator over  $n_{MC}$  replications of Monte Carlo simulations. The squared error, its mean (MSE) defined in (2.2), median, first and third empirical quartiles as well as the normalized root mean squared error (NRMSE) defined as follows

$$\text{NRMSE} \left( t, \vec{Z}_S(t) \right) := \frac{\sqrt{\text{MSE} \left( t, \vec{Z}_S(t) \right)}}{\theta_S \left( t, \vec{Z}_S(t) \right)}, \quad (3.1)$$

are used as indicators of the performances of our estimator and are calculated over a grid of  $n_t$  times  $t$  in  $[0, 1]$ .

#### 3.1 Experimental design

The Monte Carlo replications are simulated according to the model (1.1) presented in the introduction where  $\mathcal{S}$  and  $\theta_S$  are chosen as follows. On the one hand the inter-arrival times of the counting process  $M$  have the same distribution function than  $U + \varepsilon$  with  $U \sim \mathcal{E}(1/\varepsilon)$  for  $\varepsilon > 0$ , putting ourselves in the situation of Example 1. On the other hand

$$\theta_S \left( t, \vec{Z}_S(t) \right) = \lambda_0(t) \exp \left( \sin \left( \beta \cdot \vec{Z}_S(t) \right) \right), \quad (3.2)$$

where  $\lambda_0(t) = \frac{b}{a} \left( \frac{t}{a} \right)^{b-1}$  with  $a, b > 0$ ,  $\beta$  is a vector of same dimension as  $\vec{Z}_S(t)$  and  $u \cdot v$  denotes the euclidian inner product of vectors  $u$  and  $v$ . The co-process  $Z$  is simulated according to a Brownian motion.

Note that the intensity  $\theta$  is a modified version of the proportional hazards intensity function that models the dependence of our counting process on the past of the co-process  $Z$ . For  $\beta > 0$ ,  $\theta_S \left( t, \vec{Z}_S(t) \right)$  can be viewed as a stochastic perturbation



of the intensity of a Weibull process as presented in Figure 1. Remark that for  $\beta \geq 0.3$  the replications look too irregular for a kernel estimator to capture their behaviour properly. It is due to the increasing impact of the covariates which we chose to be Brownian.

As  $(N_t)_{t \in [0,1]}$  is an inhomogeneous Poisson process conditionally on  $\mathcal{Z}$  and  $\mathcal{S}$ , we can simulate its jumping times by applying the inverse function of  $\Lambda(\cdot) = \int_0^\cdot \theta_S(s, \vec{Z}_S(s)) ds$  to the jumping times of a homogeneous Poisson process with intensity 1. In our case this inverse function writes

$$\Lambda^{-1}(u) = \Lambda_0^{-1} \left( \Lambda_0(S_{j_u}) + \frac{u - \Lambda(S_{j_u})}{\exp\left(\sin\left(\beta \cdot \vec{Z}_S(t)\right)\right)} \right),$$

where  $j_u$  is such that  $\Lambda(S_{j_u}) \leq u < \Lambda(S_{j_u+1})$  and  $\Lambda_0(\cdot) = \int_0^\cdot \lambda_0(s) ds$ . This allows us to simulate the data according to our model.

We finally take

$$\begin{aligned} \mathcal{H}(u) &= \left( \frac{1}{2} - \frac{5}{8}(3u^2 - 1) \right) \mathbb{1}_{|u| \leq 1}, \\ K(u) &= (30u^2 - 36u + 9) \mathbb{1}_{0 \leq u \leq 1}, \end{aligned}$$

so that assumptions **(H5)** and **(H6)** are fulfilled.

### 3.2 Results

Figure 2 represents the theoretical intensity (solid) versus the first and third empirical quartiles (dashed and dotted) of 100 Monte Carlo replications of our estimator for  $n = 500$  and  $\beta = 0$  (top left),  $\beta = 0.1$  (top right),  $\beta = 0.3$  (bottom left) and  $\beta = 0.5$  (bottom right). The case  $\beta = 0$  is the one of the estimation of the intensity of a Weibull process. As  $\beta$  increases, the counting process  $N$  deviates from this simple case to a point where the signal is almost chaotic due to the influence of the co-process  $Z$  for  $\beta = 0.5$  (see Figure 1). As expected our estimator is less accurate for high values of  $t$  (*i.e.* high dimensionality) and quickly varying objective function (*e.g.*  $\beta = 0.5$ ). We also note an artifact for the estimation around zero. It is a well known issue with kernel estimation on the edges of the support of the objective function.

In the following we fix  $\beta = 0.1$ .

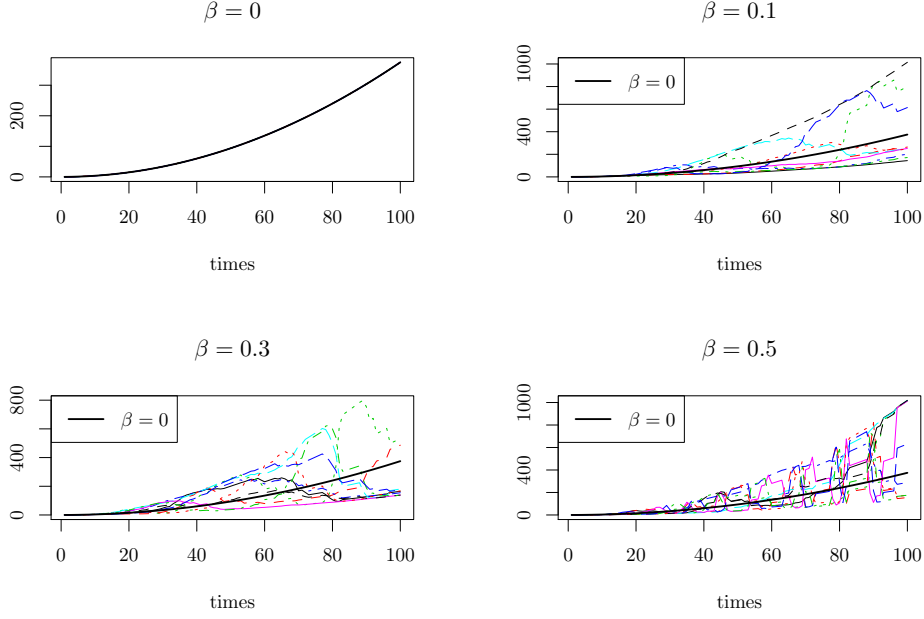


Figure 1: 10 replications of the studied intensity (3.2) for  $\beta = 0$  (top left),  $\beta = 0.1$  (top right),  $\beta = 0.3$  (bottom left) and  $\beta = 0.5$  (bottom right). The bold line represents the corresponding intensity for a Weibull process ( $\beta = 0$ ).

Figure 3 represents the median (solid line) and the first and third empirical quartiles (dashed and dotted lines) of the squared error of our estimator for 10,000 Monte Carlo replications of our estimator for  $n = 500$  (Figure 3a) and  $n = 10,000$  (Figure 3b). As expected, the results are far better for  $n = 10,000$  where the third quartile does not exceed 14,000 compared to a maximum of 50,000 for  $n = 500$ . Remark that these maxima are always attained near to  $t = 1$ . That is explained by the fact that as  $t$  increases, the dimension of the estimation problem increases.

In Table 1 the results are obtained for 10,000 Monte Carlo replications of our model for 3 different times and 5 different values of  $n$ . The dimensionality of our estimation problem increases quickly towards 59 at time  $t = 0.9$ . This shows the difficulty of the estimation for small  $n$ . We observe nevertheless an increase in performance for bigger values of  $n$ . For  $n = 10,000$  the NRMSE indicator stays below 0.2 after time  $t = 0.3$ . At this point we seem to have attained the

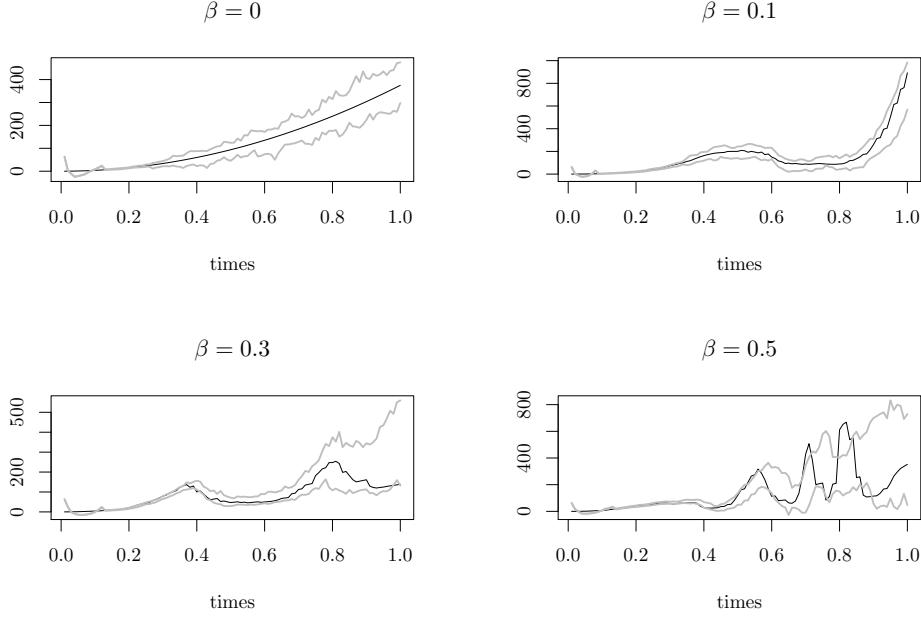
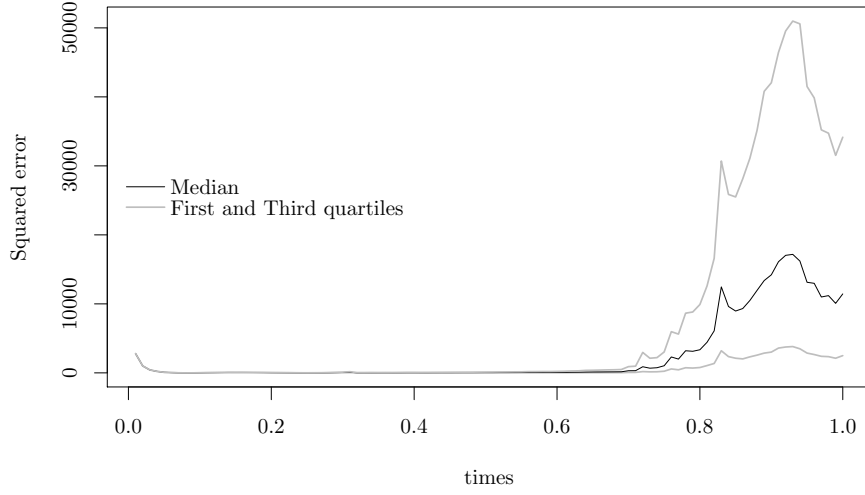


Figure 2: Objective intensity (black line) versus first and third (gray lines) empirical quartiles for 500 observations,  $n_{MC} = 100$ ,  $n_t = 100$  and  $\beta = 0$  (top left),  $\beta = 0.1$  (top right),  $\beta = 0.3$  (bottom left) and  $\beta = 0.5$  (bottom right).

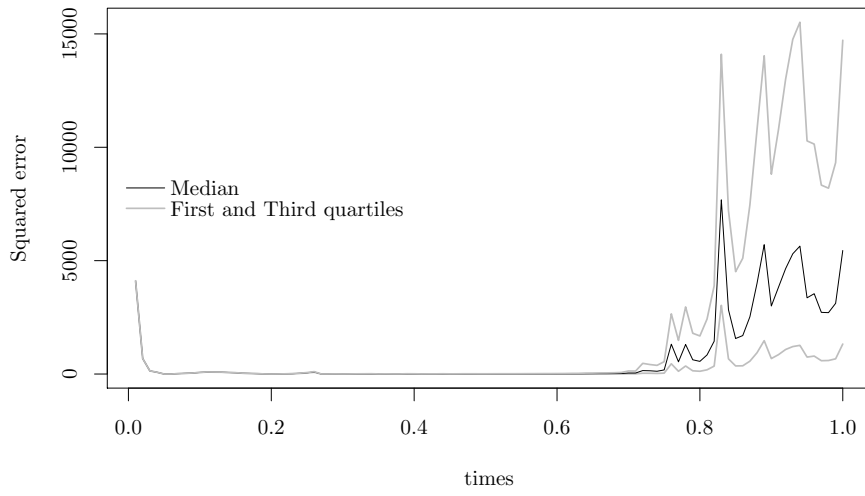
asymptotic property for the MSE described in Section 2.2.

## 4 Application to real data

We study a data set constituted of historical prices of  $n = 495$  companies as well as the crude oil prices over a period of roughly one year and two months (from 17th April, 2014 to 23rd June, 2015). The companies data are taken from the website Yahoo Finance so that every company considered composed the S&P500 index on the 23rd June, 2015. The crude oil prices are taken from the website Investing.com. The Cox process data consist of the count of the number of times when the percent returns of said companies go below a certain threshold with the counting rate depending on the stochastic dynamics of the company market capitalization. In our case the company market capitalization is represented by the action's trade volume normalized increments and is observed when the percent



(a)  $n = 500$



(b)  $n = 10,000$

Figure 3: Median, first and third empirical quartiles of the squared error of our estimator for  $n_{MC} = 10,000$ ,  $n_t = 100$ ,  $\beta = 0.1$  and (a)  $n = 500$  and (b)  $n = 10,000$ .

$n$	$t$	$M_t$	$\lambda(t)$	Estimate	MSE	NRMSE
100	0.5	33	36.21	38.15	4.7E+03	1.89
	0.7	47	110.17	97.50	1.2E+05	3.14
	0.9	59	763.96	675.75	6.5E+06	3.34
250	0.5	33	36.21	38.26	3.4E+02	0.51
	0.7	47	110.17	92.28	1.3E+03	0.33
	0.9	59	763.96	717.24	3E+06	2.27
500	0.5	33	36.21	38.37	96	0.27
	0.7	47	110.17	91.51	9.4E+02	0.28
	0.9	59	763.96	649.71	4.8E+07	9.06
1,000	0.5	33	36.21	38.03	54	0.20
	0.7	47	110.17	92.05	7.1E+02	0.24
	0.9	59	763.96	732.71	7.2E+04	0.35
10,000	0.5	33	36.21	37.84	11	0.09
	0.7	47	110.17	92.39	3.8E+02	0.18
	0.9	59	763.96	743.70	2.9E+03	0.07

Table 1: Mean value of the estimator, mean squared error (2.2) and normalized root mean squared error (3.1) for  $n_{MC} = 10,000$  and  $\beta = 0.1$ .

return of the crude oil action below another threshold. By analyzing this count, we aim to learn the financial properties of this 495 companies system.

Tables 2 and 3 present the layout of the raw data directly taken from the websites Yahoo Finance and Investing.com. We denote  $((Y_t^1)_{t \in [0,1]}, \dots, (Y_t^n)_{t \in [0,1]})$  the **Open** columns of the Yahoo Finance raw data. It represents the daily open prices of the actions of the companies. We define the percent returns as follows  $X_t^k := \frac{Y_t^k - Y_{t-1}^k}{Y_{t-1}^k}$  for  $1 \leq k \leq n$ . Denote  $(\zeta_t)_{t \in [0,1]}$  the **Open** column of the Investing.com raw data. In the same way as for the Yahoo Finance raw data it represents the daily open prices of the crude oil action. We define its percent returns by  $\xi_t := \frac{\zeta_t - \zeta_{t-1}}{\zeta_{t-1}}$ . The **Volume** columns of the Yahoo Finance data are denoted  $((W_t^1)_{t \in [0,1]}, \dots, (W_t^n)_{t \in [0,1]})$ . Its normalized increments are defined for  $1 \leq k \leq n$  by  $Z_t^k := \frac{W_t^k - W_{t-1}^k}{W_{t-1}^k}$ .

Random times  $S$  as well as the counting process  $N$  are deduced from these transformed data sets. They are defined such that  $S_1$  is the first time at which  $\xi$  goes below  $\alpha := -0.01$ ,  $S_2$  is the second time, *etc.* and the process  $N$  counts the number of times  $X$  goes below  $\beta := -0.015$ . Figure 4 represents the trajectory

Date	Open	High	Low	Close	Volume	Adj.Close
2015-06-23	39.89	39.95	39.42	39.60	2053600	39.60
2015-06-22	39.81	40.01	39.73	39.81	3901700	39.81
2015-06-19	39.80	39.94	39.49	39.49	2581000	39.49
2015-06-18	39.80	40.06	39.72	39.90	1865000	39.90
2015-06-17	39.76	39.80	39.32	39.60	1519400	39.60
2015-06-16	39.59	39.81	39.38	39.79	1422600	39.79
2015-06-15	39.63	39.63	39.25	39.52	2320100	39.52
2015-06-12	40.33	40.49	39.74	39.84	2764200	39.84
2015-06-11	40.57	40.60	40.29	40.53	1566000	40.53
2015-06-10	40.40	40.59	40.27	40.52	1787900	40.52
...						

Table 2: First 10 rows of raw data for Agilent Technologies Inc. taken from Yahoo Finance.

Date	Price	Open	High	Low	Vol.	Change
2015-06-23	61.01	60.21	61.49	59.55	336.22K	1.04%
2015-06-22	60.38	59.75	60.63	59.27	255.31K	1.29%
2015-06-19	59.97	60.88	60.93	59.24	299.89K	-1.40%
2015-06-18	60.82	60.10	61.33	59.67	171.48K	0.81%
2015-06-17	60.33	60.52	61.81	59.34	232.09K	-0.20%
2015-06-16	60.45	60.01	60.81	59.88	129.30K	0.75%
2015-06-15	60.00	60.33	60.42	59.19	128.26K	-0.66%
2015-06-12	60.40	60.92	61.06	60.18	91.96K	-1.34%
2015-06-11	61.22	61.56	61.91	60.65	150.62K	-0.97%
2015-06-10	61.82	61.00	62.22	60.88	188.78K	2.00%
...						

Table 3: First 10 rows of raw data for the crude oil action taken from Investing.com.

of  $\xi$  and gives an illustration of the construction of  $S$ . With this construction we get a total of 50 times of observation. Remark that these thresholds represent a 1% drop for the percent returns of the crude oil action and a 1.5% drop for the percent returns of the S&P500 companies action.

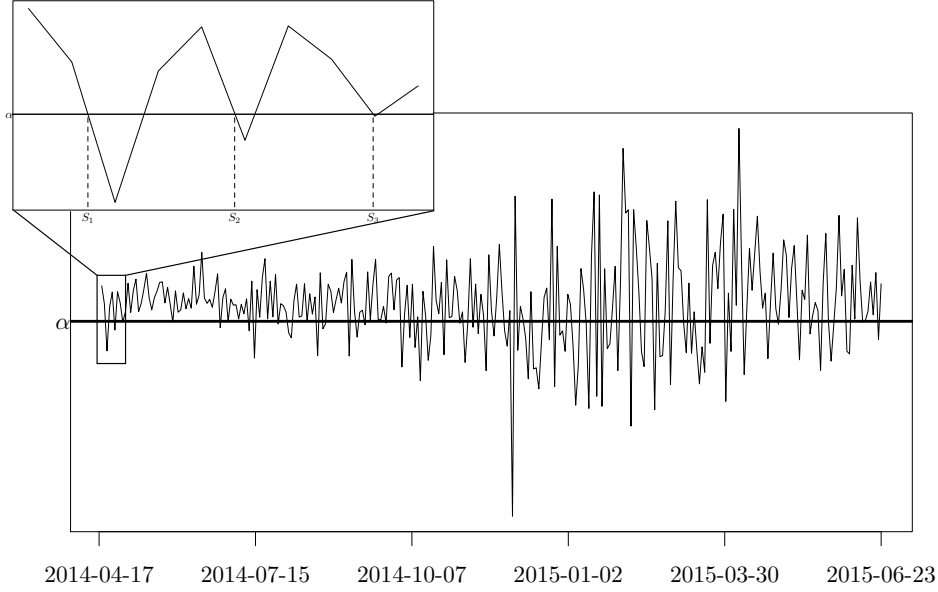


Figure 4: Crude oil action percent returns plus a zoom on a small window of time to demonstrate the construction of the random times  $S$ .

We aim to compare the inhomogeneous Poisson model with our model (1.1). To this end we compute our estimator over the time span defined by the data and for 10 chosen trajectories of the covariate process  $Z$ . The resulting estimated intensities are given in Figure 5. In most cases (7 out of 10), we estimate the same intensity as in the inhomogeneous Poisson model. In the second graph we observe that for 3 trajectories of  $Z$ , taking covariates into consideration in the model provides estimations that stays close to the inhomogeneous Poisson model at first and deviates from it after a short moment.

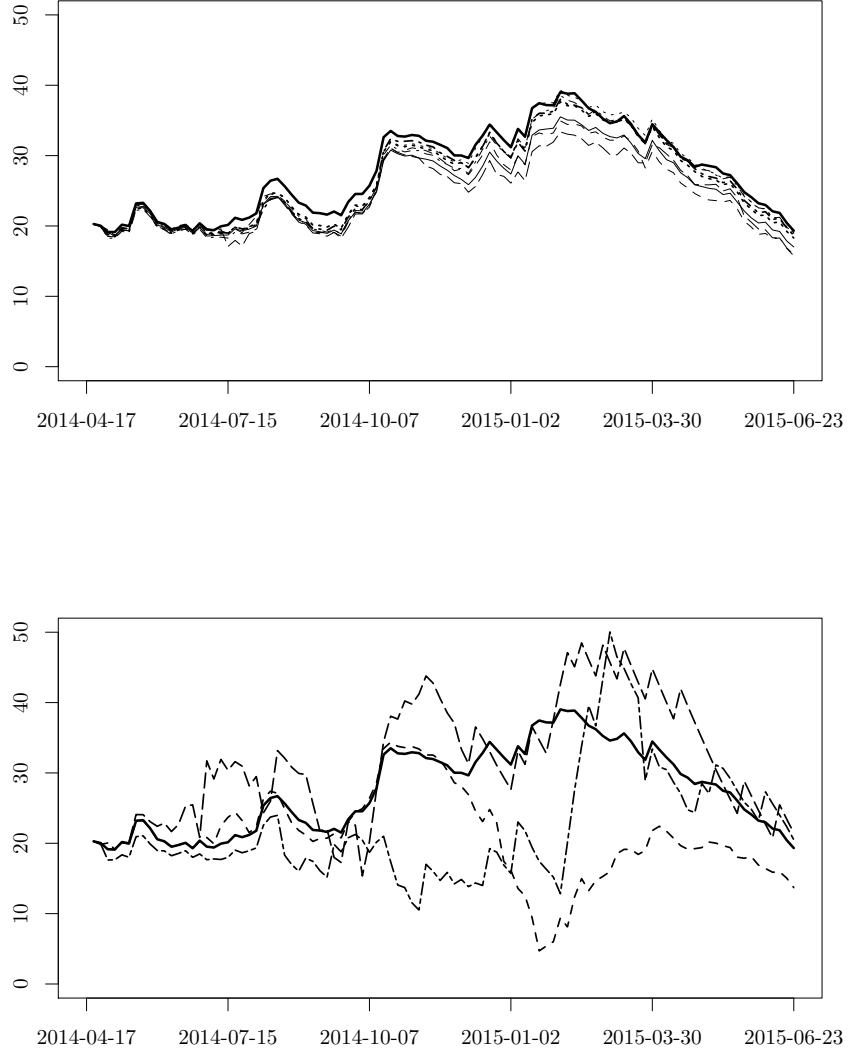


Figure 5: Estimation of the intensity function  $\lambda$  in the Cox process model (1.1) for 10 chosen trajectories of the covariate process  $Z$  compared to the estimation for an inhomogeneous Poisson model (bold solid line)



## 5 Proofs

In the sequel,  $C$  denotes a positive constant that can change of values from line to line,  $\mathbb{P}_S$ ,  $\mathbb{E}_S$  and  $\tilde{\mathbb{E}}$  respectively stand for  $\mathbb{P}(\cdot|\mathcal{S})$ ,  $\mathbb{E}(\cdot|\mathcal{S})$  and  $\mathbb{E}(\cdot|\mathcal{Z}, \mathcal{S})$ . For fixed  $t \in [0, 1]$  and  $z : [0, 1] \rightarrow \mathbb{R}$ , we define

$$\tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) = \frac{1}{n} \sum_{k=1}^n H_\eta \left( \vec{z}_S(t) - \vec{Z}_S^k(t) \right) \int_0^t K_h(t-s) \theta_S \left( s, \vec{Z}_S^k(s) \right) ds,$$

and  $\phi(t, \vec{z}_S(t)) = f(\vec{z}_S(t)) \theta_S(t, \vec{z}_S(t))$ .

### 5.1 Proof of Theorem 1

*Proof.* Define for  $k = 1, \dots, n$  the trimmed version of the Nadaraya-Watson weights as

$$\mathcal{W}_{S,\eta}^k(t, \vec{z}_S(t)) = \frac{\frac{1}{n} H_\eta \left( \vec{z}_S(t) - \vec{Z}_S^k(t) \right)}{\frac{1}{n} \sum_{l=1}^n H_\eta \left( \vec{z}_S(t) - \vec{Z}_S^l(t) \right) \vee a_n},$$

and consider

$$\bar{N}_t^k = N_t^k - \Lambda_t^k, \quad t \in [0, 1],$$

with  $\Lambda_t^k = \tilde{\mathbb{E}} N_t^k = \int_0^t \theta_S \left( s, \vec{Z}_S^k(s) \right) ds$ .

Using these notations we have

$$\tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) - \theta_S(t, \vec{z}_S(t)) = A + B, \quad (5.1)$$

where

$$\begin{aligned} A &= \sum_{k=1}^n \mathcal{W}_{S,\eta}^k(t, \vec{z}_S(t)) \int_0^t K_h(t-s) d\bar{N}_s^k, \text{ and} \\ B &= \sum_{k=1}^n \mathcal{W}_{S,\eta}^k(t, \vec{z}_S(t)) \int_0^t K_h(t-s) \theta_S \left( s, \vec{Z}_S^k(s) \right) ds - \theta_S(t, \vec{z}_S(t)). \end{aligned}$$

Decomposition (5.1) gives

$$\text{MSE}_{\mathcal{S}}(t, z) \leq 2 \left( \mathbb{E}_{\mathcal{S}} A^2 + \mathbb{E}_{\mathcal{S}} B^2 \right). \quad (5.2)$$

On the one hand, since  $((\bar{N}_t^k)_{t \in [0,1]} : k = 1, \dots, n)$  are independent trajectories of a conditional local martingale given  $\mathcal{S}$  and  $\mathcal{Z}$ ,

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} A^2 &= \mathbb{E}_{\mathcal{S}} \left( \sum_{1 \leq k \leq n} \left[ \mathcal{W}_{S,\eta}^k(t, \vec{z}_S(t)) \right]^2 \tilde{\mathbb{E}} \left[ \left( \int_0^t K_h(t-s) d\bar{N}_s^k \right)^2 \right] \right) \\ &\leq \mathbb{E}_{\mathcal{S}} \left( \sup_{k=1, \dots, n} \mathcal{W}_{S,\eta}^k(t, \vec{z}_S(t)) \tilde{\mathbb{E}} \left[ \left( \int_0^t K_h(t-s) d\bar{N}_s^1 \right)^2 \right] \right). \end{aligned}$$

As  $(\bar{N}_t^1)_{t \in [0,1]}$  is a conditional Poisson process with cumulative intensity function  $\Lambda_t^1$  given  $\mathcal{S}$  and  $\mathcal{Z}$ , the quadratic variation of  $(\bar{N}_t^1)_{t \in [0,1]}$  is the process  $(N_t^1)_{t \in [0,1]}$ . The Itô isometry hence gives

$$\tilde{\mathbb{E}} \left[ \left( \int_0^t K_h(t-s) d\bar{N}_s^1 \right)^2 \right] = \int_0^t K_h^2(t-s) \theta_S \left( t, \vec{Z}_S^1(s) \right) ds.$$

So that

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} A^2 &\leq \mathbb{E}_{\mathcal{S}} \left[ \sup_{k=1, \dots, n} \mathcal{W}_{S,\eta}^k(t, \vec{z}_S(t)) \int_0^t K_h^2(t-s) \theta_S \left( s, \vec{Z}_S^1(s) \right) ds \right] \\ &\leq \frac{\|K\|_2^2 \Theta^{M_t} \|\mathcal{H}\|_{\infty}^{dM_t}}{nh\eta^{dM_t}} \mathbb{E}_{\mathcal{S}} \left[ \frac{1}{\tilde{f}_{S,\eta}(\vec{z}_S(t))} \right] \\ &\leq \frac{\|K\|_2^2 \Theta^{M_t} \|\mathcal{H}\|_{\infty}^{dM_t}}{nh\eta^{dM_t}} \left( \frac{1}{F_0^{M_t}} + \mathbb{E}_{\mathcal{S}} \left| \frac{1}{\tilde{f}_{S,\eta}(\vec{z}_S(t))} - \frac{1}{f(\vec{z}_S(t))} \right| \right). \quad (5.3) \end{aligned}$$

Combining (5.2), (5.3), lemma 6 and lemma 7, theorem follows.  $\square$

## 5.2 Proof of Proposition 3

*Proof.* Let  $\varepsilon > 0$ , and define

$$\mathcal{A} = \left\{ \left| \tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) - \theta_S(t, \vec{z}_S(t)) \right| > \varepsilon \right\}, \text{ and}$$

$$\mathcal{B} = \left\{ \left| \tilde{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right| \leq \frac{F_0}{2} \right\}.$$

Then

$$\mathbb{P}_S(\mathcal{A}) \leq \mathbb{P}_S(\mathcal{A} \cap \mathcal{B}) + \mathbb{P}_S(\bar{\mathcal{B}}).$$

On the one hand,

$$\mathbb{E}_S \left| \tilde{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^2 \leq \mathbb{E}_S \left| \hat{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^2 + a_n^2,$$

usual properties on kernel estimation of the density [see 5] gives

$$\tilde{f}_{S,\eta}(\vec{z}_S(t)) \xrightarrow{\mathbb{L}^2} f(\vec{z}_S(t)) \text{ under } \mathbb{P}_S$$

so that  $\mathbb{P}_S(\bar{\mathcal{B}}) \xrightarrow[n \rightarrow +\infty]{} 0$ .

On the other hand,

$$\mathcal{A} \cap \mathcal{B} = \left\{ \left| \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) f(\vec{z}_S(t)) - \phi(t, \vec{z}_S(t)) \tilde{f}_{S,\eta}(\vec{z}_S(t)) \right| > \varepsilon \frac{F_0^2}{2} \right\},$$

and Lemma 8 give  $\mathbb{P}_S(\mathcal{A} \cap \mathcal{B}) \xrightarrow[n \rightarrow +\infty]{} 0$ . Combining these results with the dominated convergence theorem, proposition follows.  $\square$

## 5.3 Proof of Theorem 4

*Proof.* Define

$$A_n = \alpha_n (nh\eta^{dM_t})^{1/2} \left[ \tilde{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right], \text{ and}$$

$$B_n = \frac{\tilde{f}_{S,\eta}(\vec{z}_S(t))}{f(\vec{z}_S(t))} (nh\eta^{dM_t})^{1/2} \frac{\hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \phi(t, \vec{z}_S(t))}{\left[ \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t} \right]^{1/2}},$$

where

$$\alpha_n = \frac{\sqrt{\hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t))}}{f(\vec{z}_S(t)) \|K\|_2 \|\mathcal{H}\|_2^{dM_t}}.$$

On the one hand,

$$\alpha_n \xrightarrow{\mathbb{P}_S} \sqrt{\frac{\theta_S(t, \vec{z}_S(t))}{f(\vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t}}} \leq \sqrt{\frac{\|\theta_S\|_\infty}{F_0 \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t}}},$$

and

$$\mathbb{E}_S \left| \tilde{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^2 \leq \mathbb{E}_S \left| \hat{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^2 + a_n^2,$$

usual properties on kernel estimation of the density [see 5] gives

$$nh\eta^{dM_t} \mathbb{E}_S \left| \tilde{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^2 \xrightarrow{n \rightarrow +\infty} 0.$$

Combining these results with Slutsky lemma gives

$$A_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}_S} 0. \quad (5.4)$$

On the other hand,

$$\mathbb{E}_S \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) = (\mathcal{K}_{h,\eta} * \psi)(t, \vec{z}_S(t)),$$

with  $\mathcal{K}$  the product kernel of  $H$  and  $K$  and

$$\forall (u, y) \in [0, t] \times \mathbb{R}^{dM_t}, \psi(u, y) = f(y) \theta_S(u, \vec{y}_S(u)).$$

Applying Lemma 5 with kernel  $\mathcal{K}$  and function  $\psi$  gives

$$\mathbb{E}_S \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) \leq \phi(t, \vec{z}_S(t)) + \frac{(dM_t + 1)^2}{2} \max(h^2, \eta^2) Q_{M_t} C_{\mathcal{H},K},$$

where  $C_{\mathcal{H},K} = \max \left( \int_{\mathbb{R}} z^2 \mathcal{H}(z) dz, \int_{\mathbb{R}} u^2 K(u) du \right)$  is a finite constant since  $\mathcal{H}$  and

$K$  are compactly supported. As  $nh^5\eta^{dM_t} \xrightarrow{n \rightarrow +\infty} 0$  and  $nh\eta^{dM_t+4} \xrightarrow{n \rightarrow +\infty} 0$ ,

$$(nh\eta^{dM_t})^{1/2} \frac{\mathbb{E}_S \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \phi(t, \vec{z}_S(t))}{\left[ \phi(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t} \right]^{1/2}} \xrightarrow{n \rightarrow +\infty} 0.$$

Combining this result with Lemma 9 and Slutsky lemma gives

$$B_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \text{ under } \mathbb{P}_S. \quad (5.5)$$

As

$$(nh\eta^{dM_t})^{1/2} \frac{\tilde{\theta}_{S,h,\eta}(t, \vec{z}_S(t)) - \theta_S(t, \vec{z}_S(t))}{\left[ \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t} / \tilde{f}_{S,\eta}^2(\vec{z}_S(t)) \right]^{1/2}} = B_n - A_n,$$

using (5.4) and (5.5), we have

$$\mathbb{E}_S \left( e^{iu(B_n - A_n)} \right) \xrightarrow{n \rightarrow +\infty} e^{-\frac{u^2}{2}}.$$

Combining this result with the dominated convergence theorem, theorem follows.  $\square$

## 6 Appendix

In the sequel  $C$  denotes a positive constant under  $\mathbb{P}_S$  that can change of values from line to line and  $\text{Var}_S$  stands for  $\text{Var}(\cdot|\mathcal{S})$ . For simplicity, we may use the notations  $\tilde{f}$ ,  $\tilde{\theta}$ ,  $\hat{\phi}$  and  $\tilde{\phi}$  instead of  $\tilde{f}_{S,\eta}$ ,  $\tilde{\theta}_{S,h,\eta}$ ,  $\hat{\phi}_{S,h,\eta}$  and  $\tilde{\phi}_{S,h,\eta}$  respectively.

**Lemma 5.** *Let  $\mathbf{K}$  be a bounded, compactly supported  $d$ -variate kernel satisfying*

$$\int_{\mathbb{R}^d} \mathbf{K}(z) dz = 1 \text{ and } \int_{\mathbb{R}^d} z \mathbf{K}(z) dz = 0.$$

*Let  $\mathbf{h} = \text{diag}(h_1^2, \dots, h_d^2)$  be such that  $n^{-1} |\mathbf{h}|^{-1/2}$  and all entries of  $\mathbf{h}$  approach zero as  $n$  tends to  $+\infty$ . Also, we assume that the ratio of the largest and smallest eigenvalues of  $\mathbf{h}$  is bounded for all  $n$ .*

*Denote  $\mathbf{K}_{\mathbf{h}}(z) = |\mathbf{h}|^{-1/2} \mathbf{K}(\mathbf{h}^{-1/2} z)$ .*

Let  $f$  be a  $d$ -variate function. Also, let  $\mathfrak{D}_f(z)$  be the vector of first-order partial derivatives of  $f$  and  $\mathfrak{H}_f(z)$  be the Hessian matrix of  $f$ . Let's assume that for all  $1 \leq k \leq d$ ,

$$\forall y \in \mathbb{R}^d, \left\| \frac{\partial^2 f}{\partial x_k^2}(y) \right\|_2 \leq F_d,$$

where  $F_d$  is a positive constant depending on dimension  $d$ .

Then

$$\left| \int_{\mathbb{R}^d} \mathbf{K}_{\mathbf{h}}(x - y) f(y) dy - f(x) \right| \leq \frac{d}{2} \|\mathbf{h}\|_2 F_d \int_{\mathbb{R}^d} \|z\|_2^2 \mathbf{K}(z) dz.$$

*Proof.* By the multivariate version of Taylor's theorem with Lagrange remainder, for some  $\gamma \in (0, 1)$ ,

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \mathbf{K}_{\mathbf{h}}(x - y) f(y) dy - f(x) \right| &\leq \left| \frac{1}{2} \int_{\mathbb{R}^d} z^\top \mathbf{h}^{1/2} \mathfrak{H}_f(x - \gamma \mathbf{h}^{1/2} z) \mathbf{h}^{1/2} z \mathbf{K}(z) dz \right| \\ &\leq \frac{1}{2} \|\mathbf{h}\|_2 \sup_{y \in \mathbb{R}^d} \|\mathfrak{H}_f(y)\|_2 \int_{\mathbb{R}^d} \|z\|_2^2 \mathbf{K}(z) dz \\ &\leq \frac{d}{2} \|\mathbf{h}\|_2 F_d \int_{\mathbb{R}^d} \|z\|_2^2 \mathbf{K}(z) dz, \end{aligned}$$

as  $\int_{\mathbb{R}^d} \mathbf{K}(z) dz = 1$ , and  $\int_{\mathbb{R}^d} z \mathbf{K}(z) dz = 0$ . □

**Lemma 6.** Under the assumptions of Theorem 1, for  $p = 1$

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} \left| \frac{1}{\tilde{f}_{S,\eta}(\tilde{z}_S(t))} - \frac{1}{f(\tilde{z}_S(t))} \right|^p &\leq C_{M_t}^0 \left( (n\eta^{dM_t})^{\varepsilon-1} + \eta^2 + \frac{1}{(n\eta^{dM_t})^{1/2}} \right. \\ &\quad \left. + \frac{n\eta^{dM_t+4}}{(n\eta^{dM_t})^\varepsilon} + \frac{1}{(n\eta^{dM_t})^\varepsilon} \right), \end{aligned}$$

for  $p = 2$ ,

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} \left| \frac{1}{\tilde{f}_{S,\eta}(\tilde{z}_S(t))} - \frac{1}{f(\tilde{z}_S(t))} \right|^p &\leq C_{M_t}^1 \left( (n\eta^{dM_t})^{2\varepsilon-2} + \frac{1}{n\eta^{dM_t}} + \eta^4 + \frac{1}{(n\eta^{dM_t})^{2\varepsilon}} \right. \\ &\quad \left. + \frac{1}{(n\eta^{dM_t})^{2\varepsilon+1}} + \frac{(n\eta^{dM_t+4})^2}{(n\eta^{dM_t})^{2\varepsilon}} \right), \end{aligned}$$

and for  $p > 2$

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} \left| \frac{1}{\tilde{f}_{S,\eta}(\vec{z}_S(t))} - \frac{1}{f(\vec{z}_S(t))} \right|^p &\leq C_{M_t}^2 \left( (n\eta^{dM_t})^{p\varepsilon-p} + \frac{1}{(n\eta^{dM_t})^{p/2}} + \frac{1}{(n\eta^{dM_t})^{p-1}} \right. \\ &\quad + \eta^{2p} + \frac{1}{(n\eta^{dM_t})^{p\varepsilon}} + \frac{1}{(n\eta^{dM_t})^{p\varepsilon+p-1}} \\ &\quad \left. + \frac{(n\eta^{dM_t+4})^p}{(n\eta^{dM_t})^{p\varepsilon}} \right), \end{aligned}$$

where  $C_{M_t}^0$ ,  $C_{M_t}^1$  and  $C_{M_t}^2$  depend only on  $d$ ,  $M_t$ ,  $Q_{M_t}$ ,  $F_0$ ,  $F_\infty$ ,  $\|\mathcal{H}\|_2$ ,  $\|\mathcal{H}\|_\infty$  and  $\int_{\mathbb{R}} z^2 \mathcal{H}(z) dz$ .

*Proof.* Since

$$\frac{1}{\tilde{f}(\vec{z}_S(t))} = \frac{1}{f(\vec{z}_S(t))} - \frac{\tilde{f}(\vec{z}_S(t)) - f(\vec{z}_S(t))}{f^2(\vec{z}_S(t))} + \frac{(\tilde{f}(\vec{z}_S(t)) - f(\vec{z}_S(t)))^2}{f^2(\vec{z}_S(t))\tilde{f}(\vec{z}_S(t))},$$

we have

$$\left| \frac{1}{\tilde{f}(\vec{z}_S(t))} - \frac{1}{f(\vec{z}_S(t))} \right|^p \leq 2^{p-1} \left( \frac{|\tilde{f}(\vec{z}_S(t)) - f(\vec{z}_S(t))|^p}{F_0^{2pM_t}} + \frac{|\tilde{f}(\vec{z}_S(t)) - f(\vec{z}_S(t))|^{2p}}{F_0^{2pM_t} a_n^p} \right).$$

To control the conditional expectation of this quantity we need to control

$$\mathbb{E}_{\mathcal{S}} \left| \tilde{f}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^q \text{ for } q \geq 1.$$

We have

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} \left| \tilde{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^q &\leq \mathbb{E}_{\mathcal{S}} \left| \hat{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^q + a_n^q \\ &\leq 2^{q-1} \mathbb{E}_{\mathcal{S}} \left| \hat{f}_{S,\eta}(\vec{z}_S(t)) - \mathbb{E}_{\mathcal{S}} \hat{f}_{S,\eta}(\vec{z}_S(t)) \right|^q \\ &\quad + 2^{q-1} \left| \mathbb{E}_{\mathcal{S}} \hat{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right|^q + a_n^q. \end{aligned}$$

On the one hand, as  $\mathbb{E}_{\mathcal{S}} \hat{f}_{S,\eta}(\vec{z}_S(t)) = (H * f)(\vec{z}_S(t))$ , applying Lemma 5 with

kernel  $H$  and function  $f$  gives

$$\begin{aligned} \left| \mathbb{E}_{\mathcal{S}} \hat{f}_{S,\eta}(\vec{z}_S(t)) - f(\vec{z}_S(t)) \right| &\leq \frac{dM_t}{2} \eta^2 Q_{M_t} \int_{\mathbb{R}^{dM_t}} \|z\|_2^2 H(z) dz \\ &\leq \frac{1}{2} Q_{M_t} (dM_t \eta)^2 \int_{\mathbb{R}} z^2 \mathcal{H}(z) dz. \end{aligned}$$

On the other hand define

$$\zeta_k = H_{\eta} \left( \vec{z}_S(t) - \vec{Z}_S^k(t) \right) - \mathbb{E}_{\mathcal{S}} H_{\eta} \left( \vec{z}_S(t) - \vec{Z}_S^k(t) \right).$$

Then

$$\mathbb{E}_{\mathcal{S}} \left| \hat{f}_{S,\eta}(\vec{z}_S(t)) - \mathbb{E}_{\mathcal{S}} \hat{f}_{S,\eta}(\vec{z}_S(t)) \right|^q = \mathbb{E}_{\mathcal{S}} \left| \frac{1}{n} \sum_{k=1}^n \zeta_k \right|^q.$$

The  $\zeta_k$ 's are conditional *i.i.d.* random variable given  $\mathcal{S}$  so that for  $q \leq 2$ , Jensen inequality gives

$$\mathbb{E}_{\mathcal{S}} \left| \frac{1}{n} \sum_{k=1}^n \zeta_k \right|^q \leq \frac{F_{\infty}^{qM_t/2} \|\mathcal{H}\|_2^{q d M_t}}{(n \eta^{d M_t})^{q/2}}.$$

For  $q > 2$ , by Khintchine inequality [see 6]

$$\mathbb{E}_{\mathcal{S}} \left| \frac{1}{n} \sum_{k=1}^n \zeta_k \right|^q \leq K_q \left( \frac{F_{\infty}^{qM_t/2} \|\mathcal{H}\|_2^{q d M_t}}{(n \eta^{d M_t})^{q/2}} + \frac{2^{q-2} F_{\infty}^{M_t} (\|\mathcal{H}\|_2^2 \|\mathcal{H}\|_{\infty}^{q-2})^{d M_t}}{(n \eta^{d M_t})^{q-1}} \right),$$

where  $K_q$  is the global constant in Khintchine inequality depending only on  $q$ . Lemma follows.  $\square$

**Lemma 7.** *Under the assumptions of Theorem 1,*

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} B^2 &\leq C_{M_t}^3 \left( \frac{1}{n \eta^{d M_t}} + h^4 + \eta^4 + (n \eta^{d M_t})^{2\varepsilon-2} + \frac{1}{(n \eta^{d M_t})^{2\varepsilon}} \right. \\ &\quad + \frac{1}{(n \eta^{d M_t})^{2\varepsilon+1}} + \frac{(n \eta^{d M_t+4})^2}{(n \eta^{d M_t})^{2\varepsilon}} + \frac{1}{(n \eta^{d M_t})^2} + \frac{1}{(n \eta^{d M_t})^3} + h^8 + \eta^8 \\ &\quad \left. + (n \eta^{d M_t})^{4\varepsilon-4} + \frac{1}{(n \eta^{d M_t})^{4\varepsilon}} + \frac{1}{(n \eta^{d M_t})^{4\varepsilon+3}} + \frac{(n \eta^{d M_t+4})^4}{(n \eta^{d M_t})^{4\varepsilon}} \right). \end{aligned}$$

where  $C_{M_t}^3$  depends only on  $d, M_t, Q_{M_t}, F_0, F_{\infty}, \Theta, \|K\|_2, \|\mathcal{H}\|_2, \|\mathcal{H}\|_{\infty}$ ,



$\int_{\mathbb{R}} z^2 \mathcal{H}(z) dz$  and  $\int_{\mathbb{R}} u^2 K(u) du$ .

*Proof.* Since

$$\frac{\tilde{\phi}(t, \vec{z}_S(t))}{\tilde{f}(\vec{z}_S(t))} - \frac{\phi(t, \vec{z}_S(t))}{f(\vec{z}_S(t))} = \frac{\tilde{\phi}(t, \vec{z}_S(t)) - \phi(t, \vec{z}_S(t))}{f(\vec{z}_S(t))} + \frac{\tilde{\phi}(t, \vec{z}_S(t))}{\tilde{f}(\vec{z}_S(t))} - \frac{\tilde{\phi}(t, \vec{z}_S(t))}{f(\vec{z}_S(t))},$$

we have

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} B^2 \leq & 2 \left( \frac{\mathbb{E}_{\mathcal{S}} \left| \tilde{\phi}(t, \vec{z}_S(t)) - \phi(t, \vec{z}_S(t)) \right|^2}{F_0^{2M_t}} + (\Theta F_{\infty})^{2M_t} \mathbb{E}_{\mathcal{S}} \left| \frac{1}{\tilde{f}(\vec{z}_S(t))} - \frac{1}{f(\vec{z}_S(t))} \right|^2 \right. \\ & \left. + \frac{1}{2} \mathbb{E}_{\mathcal{S}} \left| \tilde{\phi}(t, \vec{z}_S(t)) - \phi(t, \vec{z}_S(t)) \right|^4 + \frac{1}{2} \mathbb{E}_{\mathcal{S}} \left| \frac{1}{\tilde{f}(\vec{z}_S(t))} - \frac{1}{f(\vec{z}_S(t))} \right|^4 \right). \end{aligned}$$

Note that for  $q = 2, 4$ ,

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} \left| \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \phi(t, \vec{z}_S(t)) \right|^q \leq & 2^{q-1} \left( \mathbb{E}_{\mathcal{S}} \left| \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \mathbb{E}_{\mathcal{S}} \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) \right|^q \right. \\ & \left. + \left| \mathbb{E}_{\mathcal{S}} \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \phi(t, \vec{z}_S(t)) \right|^q \right), \end{aligned}$$

where on the one hand,  $\mathbb{E}_{\mathcal{S}} \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) = (\mathcal{K}_{h,\eta} * \psi)(t, \vec{z}_S(t))$  with  $\mathcal{K}$  the product kernel of  $H$  and  $K$  and

$$\forall (u, y) \in [0, t] \times \mathbb{R}^{dM_t}, \psi(u, y) = f(y) \theta_S(u, \vec{y}_S(u)).$$

Applying Lemma 5 with kernel  $\mathcal{K}$  and function  $\psi$  gives

$$\left| \mathbb{E}_{\mathcal{S}} \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \psi(t, \vec{z}_S(t)) \right| \leq \frac{1}{2} (dM_t + 1)^2 C_{\mathcal{H},K} Q_{M_t} \max(\eta^2, h^2),$$

where  $\psi(t, \vec{z}_S(t)) = \phi(t, \vec{z}_S(t))$  and  $C_{\mathcal{H},K} = \max(\int_{\mathbb{R}} z^2 \mathcal{H}(z) dz, \int_{\mathbb{R}} u^2 K(u) du)$  is a finite constant since  $\mathcal{H}$  and  $K$  are compactly supported. On the other hand define

$$\xi_k = H_{\eta} \left( \vec{z}_S(t) - \vec{Z}_S^k(t) \right) \int_0^t K_h(t-s) \theta_S \left( s, \vec{Z}_S^k(s) \right) ds - \mathbb{E}_{\mathcal{S}} \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)).$$

Then

$$\mathbb{E}_{\mathcal{S}} \left| \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \mathbb{E}_{\mathcal{S}} \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) \right|^q = \mathbb{E}_{\mathcal{S}} \left| \frac{1}{n} \sum_{k=1}^n \xi_k \right|^q.$$

The  $\xi_k$  are conditional *i.i.d.* centred random variables given  $\mathcal{S}$  so that for  $q \leq 2$ , by Jensen inequality

$$\mathbb{E}_{\mathcal{S}} \left| \frac{1}{n} \sum_{k=1}^n \xi_k \right|^q \leq \frac{F_{\infty}^{qM_t/2} \Theta^{qM_t} \|\mathcal{H}\|_2^{q d M_t}}{(n \eta^{d M_t})^{q/2}}.$$

For  $q > 2$ , Khintchine inequality gives

$$\mathbb{E}_{\mathcal{S}} \left| \frac{1}{n} \sum_{k=1}^n \xi_k \right|^q \leq K_q \left( \frac{F_{\infty}^{qM_t/2} \Theta^{qM_t} \|\mathcal{H}\|_2^{q d M_t}}{(n \eta^{d M_t})^{q/2}} + \frac{2^{q-2} F_{\infty}^{M_t} \Theta^{qM_t} (\|\mathcal{H}\|_2^2 \|\mathcal{H}\|_{\infty}^{q-2})^{d M_t}}{(n \eta^{d M_t})^{q-1}} \right).$$

Using Lemma 6 to control  $\mathbb{E}_{\mathcal{S}} \left| \frac{1}{\tilde{f}(\vec{z}_S(t))} - \frac{1}{f(\vec{z}_S(t))} \right|^p$  for  $p = 2, 4$ , lemma follows.  $\square$

**Lemma 8.** *Under the assumptions of Proposition 3, almost surely*

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} \hat{\phi}_{s,h,\eta}(t, \vec{z}_S(t)) &\xrightarrow{n \rightarrow +\infty} \phi(t, \vec{z}_S(t)) \text{ and} \\ n h \eta^{d M_t} \text{Var}_{\mathcal{S}} \hat{\phi}_{s,h,\eta}(t, \vec{z}_S(t)) &\xrightarrow{n \rightarrow +\infty} \phi(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2 d M_t}. \end{aligned}$$

*Proof.* We have

$$\mathbb{E}_{\mathcal{S}} \hat{\phi}_{s,h,\eta}(t, \vec{z}_S(t)) = (\mathcal{K}_{h,\eta} * \psi)(t, \vec{z}_S(t)),$$

with  $\mathcal{K}$  the product kernel of  $H$  and  $K$  and

$$\forall (u, y) \in [0, t] \times \mathbb{R}^{d M_t}, \psi(u, y) = f(y) \theta_S(u, \vec{y}_S(u)).$$

Applying Lemma I.4 from [5], with kernel  $\mathcal{K}$  and function  $\psi$  gives

$$\mathbb{E}_{\mathcal{S}} \tilde{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) \xrightarrow{n \rightarrow +\infty} \psi(t, \vec{z}_S(t)) = \phi(t, \vec{z}_S(t)). \quad (6.1)$$

Also,

$$\begin{aligned}\mathbb{E}_S \hat{\phi}_{S,h,\eta}^2(t, \vec{z}_S(t)) &= \frac{1}{n} \mathbb{E}_S H_\eta^2 \left( \vec{z}_S(t) - \vec{Z}_S^1(t) \right) \tilde{\mathbb{E}} \left( \int_0^t K_h(t-s) dN_s^1 \right)^2 \\ &\quad + \frac{n-1}{n} \mathbb{E}_S^2 \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)).\end{aligned}$$

As  $N$  is a conditional Poisson process given  $\mathcal{S}$  and  $\mathcal{Z}$ ,

$$\begin{aligned}\tilde{\mathbb{E}} \left( \int_0^t K_h(t-s) dN_s^1 \right)^2 &= \left( \int_0^t K_h(t-s) \theta_S \left( s, \vec{Z}_S^1(s) \right) ds \right)^2 \\ &\quad + \int_0^t K_h^2(t-s) \theta_S \left( s, \vec{Z}_S(s)^1 \right) ds.\end{aligned}$$

Define

$$\mathcal{K}_{h,\eta}^1(u, v, y) = K_h(u) K_h(v) H_\eta^2(y), \quad \mathcal{K}_{h,\eta}^2(u, y) = K_h^2(u) H_\eta^2(y)$$

and  $\psi_1$  a  $(dM_t + 2)$ -variate function such that for all  $(u, v, y) \in [0, t]^2 \times \mathbb{R}^{dM_t}$ ,  $\psi_1(u, v, y) = f(y) \theta_S(u, \vec{y}_S(u)) \theta_S(v, \vec{y}_S(v))$ . Then

$$\begin{aligned}\text{Var}_S \hat{\phi}_{S,h,\eta}^2(t, \vec{z}_S(t)) &= \frac{1}{nh\eta^{dM_t}} (\mathcal{K}_{h,\eta}^2 * \psi) (t, \vec{z}_S(t)) \\ &\quad + \frac{1}{n\eta^{dM_t}} (\mathcal{K}_{h,\eta}^1 * \psi_1) (t, t, \vec{z}_S(t)) - \frac{1}{n} \mathbb{E}_S^2 \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)).\end{aligned}$$

Applying Lemma I.4 from [5], with kernel  $\mathcal{K}^1 / \|\mathcal{H}\|_2^{2dM_t}$  and function  $\psi^1$ , kernel  $\mathcal{K}^2 / \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t}$  and function  $\psi$  and using (6.1) gives

$$nh\eta^{dM_t} \text{Var}_S \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) \xrightarrow{n \rightarrow +\infty} \phi(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t}.$$

□

**Lemma 9.** *Under the assumptions of Theorem 4,*

$$(nh\eta^{dM_t})^{1/2} \frac{\hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \mathbb{E}_S \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t))}{\left( \phi(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t} \right)^{1/2}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

*Proof.* Denote

$$\hat{L}_k = (nh\eta^{dM_t})^{1/2} \frac{H_\eta \left( \vec{z}_S(t) - \vec{Z}_S^k(t) \right) \int_0^t K_h(t-s) dN_S^k - \mathbb{E}_S \hat{\phi}_{S,\eta}(t, \vec{z}_S(t))}{n \left( \phi(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t} \right)^{1/2}},$$

and

$$\hat{L} = (nh\eta^{dM_t})^{1/2} \frac{\hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) - \mathbb{E}_S \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t))}{\left( \phi(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t} \right)^{1/2}},$$

so that  $\hat{L} = \sum_{k=1}^n \hat{L}_k$ . The variables  $(\hat{L}_1, \hat{L}_2, \dots)$  are conditional independent random variables given  $\mathcal{S}$  and

$$\forall k = 1, \dots, n, \mathbb{E}_S \hat{L}_k = 0 \text{ and } \text{Var}_S \hat{L} = \sum_{k=1}^n \text{Var}_S \hat{L}_k.$$

To conclude it remains to check that the Lyapunov condition

$$(\text{Var}_S \hat{L})^{-(2+\delta)/2} n \mathbb{E}_S \left| \hat{L}_1 \right|^{2+\delta} \rightarrow 0, \quad (6.2)$$

is satisfied for some  $\delta > 0$ . Remark that

$$\text{Var}_S \hat{L} = \frac{nh\eta^{dM_t}}{\phi(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t}} \text{Var}_S \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)),$$

and Lemma 8 gives

$$nh\eta^{dM_t} \text{Var}_S \hat{\phi}_{s,h,\eta}(t, \vec{z}_S(t)) \xrightarrow{n \rightarrow +\infty} \phi(t, \vec{z}_S(t)) \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t},$$

so that  $\text{Var}_S \hat{L} \rightarrow 1$  as  $n \rightarrow +\infty$ . It suffices to show that  $n \mathbb{E}_S (|\hat{L}_1|^{2+\delta})$  tends to 0 for some  $\delta > 0$  to get the Lyapunov condition (6.2) satisfied. Let us take  $\delta = 2$ ,

then

$$\begin{aligned} \mathbb{E}_{\mathcal{S}} \left| \hat{L}_1 \right|^4 &\leq \left( \frac{nh\eta^{dM_t}}{\phi \|K\|_2^2 \|\mathcal{H}\|_2^{2dM_t}} \right)^2 \frac{9}{n^4} \left[ \mathbb{E}_{\mathcal{S}} \left| H_{\eta} \left( \vec{z}_S(t) - \vec{Z}_S^1(t) \right) \int_0^t K_h(t-s) d\bar{N}_s^1 \right|^4 \right. \\ &\quad + \mathbb{E}_{\mathcal{S}} \left| H_{\eta} \left( \vec{z}_S(t) - \vec{Z}_S^1(t) \right) \int_0^t K_h(t-s) \theta_S \left( s, \vec{Z}_S^1(s) \right) ds \right|^4 \\ &\quad \left. + \mathbb{E}_{\mathcal{S}}^4 \hat{\phi}_{S,h,\eta}(t, \vec{z}_S(t)) \right]. \end{aligned}$$

Basic martingale properties as well as the Burkholder-Davis-Gundy inequality give

$$n\mathbb{E}_{\mathcal{S}} \left| \hat{L}_1 \right|^4 \leq C \frac{h^2 \eta^{2dM_t}}{n} \left( \frac{1}{\eta^{dM_t}} + \frac{1}{\eta^{3dM_t}} + h^8 + \eta^8 + 1 \right),$$

where the constant  $C$  depends only on  $\|\mathcal{H}\|_2$ ,  $\|K\|_2$ ,  $\int_{\mathbb{R}} z^2 \mathcal{H}(z) dz$ ,  $\int_{\mathbb{R}} z^2 K(z) dz$ ,  $\|\theta_S\|_{\infty}$ ,  $F_{\infty}$ ,  $Q_{M_t}$  and  $dM_t$ . We can conclude that  $n\mathbb{E}_{\mathcal{S}}(\hat{L}_1^4) \rightarrow 0$  as  $n \rightarrow +\infty$ . Lemma follows.  $\square$

## Acknowledgement

## References

- [1] S. Asmussen and H. Albrecher. *Ruin probabilities*, volume 14. World Scientific Publishing Company, 2010.
- [2] W. Bialek, F. Rieke, R. De Ruyter van Steveninck, and D. Warland. Reading a neural code. *Science*, 252(5014):1854–1857, 1991.
- [3] G. Biau, F. Cérou, and A. Guyader. Rates of convergence of the functional-nearest neighbor estimate. *Information Theory, IEEE Transactions on*, 56(4):2034–2040, 2010.
- [4] Peter J Bickel. On adaptive estimation. *The Annals of Statistics*, pages 647–671, 1982.
- [5] D. Bosq and J. P. Lecoutre. *Théorie de l'estimation fonctionnelle*, volume 21. Economica, 1987.

- [6] J. Bretagnolle and C. Huber. Estimation des densités: risque minimax. *Probability Theory and Related Fields*, 47(2):119–137, 1979.
- [7] R. Brette. Generation of correlated spike trains. *Neural computation*, 21(1): 188–215, 2008.
- [8] L. Györfi, M. Kohler, A. Krzyżak, and H. Walk. *A distribution-free theory of nonparametric regression*. Springer Science & Business Media, 2002.
- [9] S. C. Kou, X. Sunney Xie, and J. S. Liu. Bayesian analysis of single-molecule experimental data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3):469–506, 2005.
- [10] M. Krumin and S. Shoham. Generation of spike trains with controlled auto- and cross-correlation functions. *Neural Computation*, 21(6):1642–1664, 2009.
- [11] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1):125–144, 1976.
- [12] Y. Ogata. Statistical models for earthquakes occurrences and residual analysis for point processes. *Journal of the Royal Statistical Society*, 44:102–107, 1988.
- [13] F. O’Sullivan. Nonparametric estimation in the cox model. *The Annals of Statistics*, pages 124–145, 1993.
- [14] T. Zhang and S. C. Kou. Nonparametric inference of doubly stochastic poisson process data via the kernel method. *The annals of applied statistics*, 4(4):1913, 2010.